Stages of User Engagement on Social Commerce Platforms : Analysis with the Navigational Clickstream Data

Kumar, Ashish

2019-04-03

Kumar , A , Salo , J & Li , H 2019 , ' Stages of User Engagement on Social Commerce Platforms : Analysis with the Navigational Clickstream Data ' , International journal of electronic commerce , vol. 23 , no. 2 , pp. 179-211 . https://doi.org/10.1080/10864415.2018.1564550

http://hdl.handle.net/10138/321657
https://doi.org/10.1080/10864415.2018.1564550

acceptedVersion

Stages of User Engagement on Social Commerce Platforms: Analysis with the Navigational Clickstream Data

Ashish Kumar
Assistant Professor of Marketing
School of Business, Aalto University
P.O. Box 21230, Lapuankatu 2
FI-00076 Aalto, Finland
Phone: +358 504 311 813
Email: ashish.kumar@aalto.fi

Jari Salo
Professor of Marketing
Department of Economics and Management, University of Helsinki
P.O. BOX 27, Latokartanonkaari 5
FI-00014, Helsinki, Finland
Phone: +358 504 489 984
Email: jari.salo@helsinki.fi

Hongxiu Li
Assistant Professor of Business Data Analytics
Department of Industrial and Information Management, Tampere University of Technology
P.O. Box 527, Korkeakoulunkatu 10
FI-33101, Tampere, Finland
Phone: +358 503 722 471
Email: hongxiu.li@tut.fi

Author Biographies

Ashish Kumar is an assistant professor at the Department of Marketing, Aalto University School of Business. His research interests include digital marketing, social media marketing, multichannel retailing and health marketing. He received his Ph.D. in marketing from the School of Management and an MA in economics from the College of Arts and Sciences, respectively, from the State University of New York at Buffalo. His research has been published in journals such as the *Journal of Marketing*, *Information Systems Research*, *Journal of Retailing*, *Journal of the Association for Consumer Research*, *Journal of Business Research*, *European Journal of Marketing*, *Journal of Marketing Communication*, *Journal of Retailing and Consumer Services*, *Corporate Reputation Review*, and *Applied Economics Letters.*

Jari Salo holds D.Sc. (Econ. & Bus. Adm.) from the University of Oulu, Oulu Business School (AACSB), Finland. He is a Professor of Marketing at the University of Helsinki, Finland. Previously, he was Professor of Marketing and John Garlick Chair in Business Science at the University of Cape Town, South Africa. He has over 160 scientific publications including books. He has been teaching over 20 different courses in the area of marketing and management for doctoral students, master and bachelor level students in Finland, Greece, Australia, South Africa and Italy. He has also been involved in executive education for over 15 years in Finland specializing in digital marketing and digital business.

Hongxiu Li is an assistant professor at the Department of Industrial and Information Management, Tampere University of Technology, Finland. Her expertise and research interests cover the areas of IS adoption and post-adoption behavior, consumer behavior, e-services, social media, eWOM and big data. She has more than 50 research publications in outlets such as *Information Systems Journal*, *European Journal of Information Systems*, *Computers in Human Behavior*, *Computer & Education*, *Decision Support Systems*, *Tourism Management*, *Information & Management*, and *PLOS ONE*.

Stages of User Engagement on Social Commerce Platforms: Analysis with the Navigational Clickstream Data

Abstract

Social commerce platforms have gained prominence in e-commerce, as social media has become an integral part of users' online activities. Therefore, firms have been either developing or utilizing social commerce platforms to increase user engagement by adding social shopping facility onto their electronic commerce platforms. However, managing user engagement and user interaction become complex when e-commerce platforms are transformed into social commerce platform. In this study, we operationalize four distinct stages of social commerce platform namely social identification, social interaction, social shopping and transaction based on salience theory. Using clickstream data, we empirically measure user engagement in these four states by modeling users' incidence and time spent. Drawing from the PageRank algorithm, we capture the importance of ranking and distance on user engagement. The model also accounts for the effects of situational variables such as weekend, holiday, time of the day, and user characteristics such as gender, and social media setting. Our results suggest ranking and distance have significant effects on users' incidence as well as time spent on social commerce platform. The insights from this study can be helpful in designing the social commerce platform effectively using only the customers' path navigational clickstream data from the parent social commerce platform.

KEY WORDS AND PHRASES: social commerce platform; online shopping; PageRank algorithm; Dijkstra's shortest path algorithm; online social communities; multivariate type-2 Tobit; hierarchical Bayesian method; clickstream data.

In recent years, the concept of user community has become the focus of many firms, and it has undergone a series of overhauls. For example, what started as brand communities [1, 51, 68] soon evolved into social commerce platform [6, 37, 47, 79] built on electronic commerce [80]. In 2017, with more than 70 percent of the US internet users indulging in social networks with an average of 1.15 hours spent on social media platforms [23], social commerce platforms are touted as next big frontier for retail engagement by industry [13, 17, 72] as well academia [18, 76]. The importance of social media in influencing consumers' purchase decision has seen earnings of top 500 retailers from social shopping at $6.5 billion in 2017, a 24 percent increase from the previous year [9]. Thus, by allowing users not only to consume products but also to converse, advocate, and bond with them [22, 55, 70], these social commence platforms are adding tremendous value to retailers. However, the complexity of the platforms increases, as they have to manage both commercial and social activities to provide a superior shopping experience [11, 44].

In general, social commerce platforms constitute three essential elements: social media technologies, social interaction, and commercial activities [28, 45, 79, 80]. Firms have adopted these critical elements into their social shopping strategy in three major configurations. First, they are using public social networking websites to add commercial features such as advertisements and branding [8]. Second, e-commerce firms are adding social networking capabilities to their websites to serve their customers' social needs. The third configuration is a full-blown social commerce platform or site to serve the commercial and social needs of specific interest groups.

In this study, we focus on the third configuration of the social commerce platform. The success of such social commerce platforms depends on primarily two factors: outcome of social interactions (such as word of mouth, trusted advice from experts and friends), and outcome of commercial activities (such as users' purchase, and product adoption) [41, 80]. Given the complexities involved in managing both social and commercial activities, the design of the social commerce platform is critical. The specific design problems relate to promoting users' social and commercial activities to enhance user engagement given there are no standards across such platforms, and how to make sense of user activities given an enormous amount of structured and unstructured data are generated per user in a tiny amount of time [5, 35, 36].

In this research, we address the issue of user engagement on social commerce platforms. Our study answers following research question: *how do users' social and commercial activities affect user engagement on social commerce platform?* In this regard, first, we propose a conceptual framework to establish four distinct stages of social commerce platform namely social identification, social interaction, social shopping, and transaction. Second, we empirically examine user engagement in such platform using clickstream data. To achieve our goal, we employ multiple methods from computer science (e.g., PageRank and Dijkstra's algorithm) [20, 60] to operationalize relevant variables (such as rank and distance), and econometrics (Tobit model, Bayesian methods, and MCMC) to model users' incidence and time spent in four stages of social commerce platforms. Conceptual framework together with empirical analysis enables us to get insights into user engagement on such platforms. Our results suggest that ranking and distance have significant effects on users' incidence as well as time spent on social commerce platform.

Literature Review

The objective of social commerce platform is to turn purchases into the conversation and vice versa [6, 80]. Following Liang and Turban [45], we define social commerce platforms as services that combine e-commerce activities and transactions with the social media environment. Such platforms are widely prevalent in fashion industry [55, 71], which gradually spread into a wide range of other commercial areas providing a common platform to shoppers, designers, stylists, experts, retailers, and manufacturers. Various participants on these platforms share their ideas often providing customers with customized shopping experience based on their profiles, preferences and the wisdom of crowds and their friends [78]. The core idea of such platforms is to provide a unique social shopping experience. For example, Amazon has acquired Quorus to add social experience to its online shopping platform, and eBay has launched eBay Social Shopping in the UK to harness the power of social media where product recommendations from various social networks are pulled together to enhance users' shopping experience [73].

Various studies have looked at social commerce platforms from social, commercial, design, and value perspectives. From a social standpoint, studies look into the impact of users' online social participation on their behavior [2, 65], measuring online word of mouth [25, 26], diffusion of user-generated contents [33, 49], and the impact of information technologies on online knowledge creation [48]. From the commercial perspective, studies look into users' shopping activities through the lens of social media [4, 6, 24, 26, 32, 74, 75] and multi-channel coordination [63]. These studies provide insights into how social activities influence purchase behavior. From the design perspective, studies provide insights into the effects of interactive decision aids [30], predicting internet buying behavior [11], and turning website visitors into customers [31, 64, 77]. From the value perspective, studies have looked into measuring customer experience in an online shopping environment [56] and developing marketing models for e-business [49, 80]. However, many of these studies use either social or commercial activities, and use data that do not capture the granular level of user activities on social commerce platform. Traditionally, studies have used clickstream data to uncover users' online navigational patterns [3]. Clickstream data are electronic records of users' online activities that trace their navigational path at a granular level. The advantage of clickstream data is its detail and wide coverage of online activities collected in user's natural environment, and the insights gained from clickstream data can be extended to offline setting also [11]. Clickstream data have been used in a wide variety of setting to understand online behavior such as browsing and path analysis [11, 54, 57, 62]. They have also been used to understand users' purchase behavior [69], online search behavior [37], internet-portal choice [28], page views [19], exposure to banner ad [66], and product viewing [53]. In Appendix A, we highlight the use of clickstream data to understand online user behavior.

In this study, we fill the gap in the literature by exploring user engagement in the social commerce platform using clickstream data. In this regard, we propose a conceptual framework, operationalize meaningful metrics, and develop an econometric model [12]. Our multimethod approach allows analyzing clickstream data efficiently to uncover an underlying pattern of user engagement on the social commerce platform. In Appendix B, we present the contribution of our study vis-à-vis the extant studies on social shopping phenomenon.

Conceptual Framework

In Figure 1, we present our conceptual framework. The framework comprises of three parts: first, conceptualization of social commerce platform, second, defining performance metrics, and third, factors that affect these performance metrics.

---------- *Insert Figure 1 about here* ----------

Conceptualization of Social Commerce Platforms

Yadav et al. [79] define social commerce as exchange-related activities that occur in the computer-mediated social environment. These exchange-related activities constitute need recognition, pre-purchase, purchase and post-purchase. Thus, social commerce platforms derive economic values by combining user' social activities with their commercial activities [80]. Hence, the core idea of such platforms is built upon the notion of users being social, and they indulge in consumption as much as they prefer to converse about it, which makes it different from e-commerce. Therefore, on such platforms, social media is central to shopping that provides users with various social activities that augment their commercial activities. In this regard, Huang and Benyoucef [34] identify two vital components of social commerce: exchange-related activities and computer-mediated social environment. The former component constitutes stages of consumer purchase decision-making while the latter component constitutes social connection, interactions, and communications. Thus, following Liang and Turban [45], Huang and Benyoucef [34], and Yadav et al. [79], we categorize users' activities into social and commercial on these platforms.

Social Activities. The social activities of social commerce platforms are a critical part of the computer-mediated social environment. These social activities facilitate sociability. Sociability aspects of social commerce platform offer rich social experience to consumers in which they can connect with friends and groups with similar interests, share their consumption experiences, and communicate with others. In this regard, Huang and Benyoucef [34] identify participation, conversation, and community as three key design elements that facilitate sociability. Participation supports consumers becoming part of the social commerce platform enabling them to generate, edit, share and disseminate information [9]. Conversation helps consumers in establishing a relationship with participating consumers on social commerce platform. Finally, community element enables networking elements and collaboration. Drawing from these studies [9, 34, 79, 80], we classify two primary states of consumers' social activities on social commerce platforms: social identification and social interaction. The participation in the social commerce platform takes place via activities related to the social identification. It mainly constitutes users signing up or becoming a member of a social commerce platform. Notice that this is the first precursor towards users' ultimate indulgence into any other social or commercial activities. The conversation and building community on such platforms take place via activities related to social interactions.

Commercial Activities. We note that the exchange-related activities are another important component of the social commerce platform [79]. Notably, exchange-related activities mostly constitute consumers' purchase decision-making process such as selection and ordering of products. Lian and Lai [47] propose five stages of consumer decision making processes in an online environment that include need recognition, searching, evaluation, purchase, and post-purchase. Online store design considerations widely take these processes into account [7]. Thus, based on these past studies [7, 47, 79] we categorize commercial activities into two primary states: social shopping and transaction. Social shopping activities constitute viewing products, adding or removing them into a shopping basket. *Transaction* activities constitute checking out and finally placing the order [25].

Performance Metrics

The social aspects of social commerce platforms make the issue of user engagement central to them [43, 34, 45, 59]. In this regard, we capture the user engagement on such platform using two metrics namely incidence and time spent. Incidence captures whether a user visits a particular state of the social commerce platform (social identification, social interaction, social shopping, and transaction) or not. Thus, it is a binary construct that takes value 1 for a given state of the social commerce platform if a user visits that state, otherwise it is 0. Time spent is a continuous construct that captures the total amount of time spent in a particular state by a user having visited that state. We observe a user's time spent on a state conditional on his/her incidence to that state. Past studies have used clickstream data to model user engagement using incidence and time spent using type-II Tobit model [11, 12] as well as state transition [54] and purchase behavior [57].

Factors Affecting User Engagement on Social Commerce Platforms

User Online Activity. Drawing from the salience theory[1] [73] from psychology and PageRank algorithm developed in computer science [39] and used in various organization studies [59], social media [15], marketing [14, 39], and business decision making [1], we propose that *rank* and *distance* metrics operationalized through users' online activities on social commerce platform can explain their engagement on such platform. We describe the details of how to operationalize these variables in the Methodology section.

User Characteristics. Heterogeneity plays an important role in explaining differences in online behavior across users [38]. Drawing from heterogeneity concept, we propose that certain user characteristics such as user' demographics (e.g., age, gender, and race) and psychographics (e.g., technology and online risk perception) will have a differential impact on user engagement on social commerce platform. These user level differences can be useful in segmenting the users of social commerce platforms. Therefore, in our model, we account for the effects of user characteristics that could influence their engagement on such platforms.

Situational Factors. Various situational factors affect user engagement in the online environment [61]. Therefore, we expect situational factors to affect user engagement in social commerce platform. Hence, we include situational factors such as time, weekend, and holidays in our framework as covariates of user engagement in social commerce platform.

The Data

In this section, we describe the clickstream data used in this study. Our dataset comes from a panel of online users who join a social commerce platform provided by Digital Foodie Ltd.[2] in Finland

[1] We divulge some relevant details regarding salience theory in the Appendix C.
[2] http://www.digitalfoodie.com/

to discuss and purchase food recipes. The company was founded in December 2009 with its first presence in UK market. It maintains a multiplatform service Foodie.fm[3] that provide on-demand grocery solutions. Briefly, the company offers a social media platform based on users' grocery shopping needs where in addition to engaging and interacting with topics related to food recipes, users can order the ingredients from the participating grocery store chains. The company's product suite can run on multiple platforms such as smartphones, tablets, and personal computers. Next, we describe the user activities on this platform.

States of Social Commence Platform Based on User Activities

The Foodie website provides social commerce platform accessible via multiple devices (such as mobile, tablet, and personal computer) to users with grocery needs. On this platform, users can perform various social and commercial activities. The social activities on the platform revolve around engagement and interaction among users on the topics related to various food recipes. Commercial activities constitute mostly shopping related processes.[4] Thus, these social and commercial activities together help in identifying the four states of the platform. On the Foodie platform, users can perform 22 difference actions. We categorize these actions into four distinct states (social activities - identification and interaction; commercial activities – social shopping and transaction). We present this categorization in Table 1. In the clickstream data, action code is used instead of action name to save space.

---------- *Insert Table 1 about here* ----------

[3] https://www.foodie.fi/ - This platform is in operation in Finland.
[4] According to the company, "*With Foodie Recipes you can seamlessly link your recipe bank to product purchases and store assortments. Consumers can create and share their own favorite recipes, which Foodie will automatically link to your assortment. Foodie Recipes also includes recipe recommendations tailored to consumers' individual taste profiles and preferences.*"

Three social activities related to social identification namely, creating a user account, returning or logging in at different points in time, and inviting friends to join the network. In our research context, there are eight social activities related to a social interaction where users can comment, vote, like, and unlike products and recipes. Furthermore, in our research context, there are eight commercial activities related to shopping where users can view products and recipes, and they can add or remove these items from their shopping basket. Finally, three user activities are related to checking out process. They are categorized as transaction-related shopping activities.

Sample Selection and Descriptive Analysis

Sample Selection. The duration of our dataset is six months from April 16, 2012, to October 16, 2012. We coin following terms that describe users' browsing behavior. We define an *action* as any of the 22 activities that a user can perform on this platform. A *session* is defined as a sustained period of being active on the platform. If a user has not taken any further *action* for 15 minutes or more, or there is new user action of *ReturningUser*[5] meaning the user has logged in, we assume that the session has ended, and the next new *action* marks the beginning of a new *session*. We define a *state* as categories of user actions divided into social and commercial activities. Following our framework, we define four states namely social identification, social interaction, social shopping, and transaction; the former two relate to social activities, whereas, the latter two relate to commercial activities. Finally, *active period* is defined as the total number of days a user has been to the platform during the study period. We select 500 users who have at least one session during the study period with information on their demographics and online privacy settings. We collected the demographic information using an online survey. In Table 2, we present the total

[5] Note that users accessing the social shopping website especially via mobile or tablets are always logged in. Therefore, we cannot distinguish beginning of new session for such users as *ReturningUser* action relates to users' logging in process.

counts of different actions taken by the sampled users for the duration of the study period. In Table 3, we present some of the key data descriptive.

---------- *Insert Table 2 and Table 3 about here* ----------

Social Shopping Path. These 22 different actions categorized into social (social identification and social interaction) and commercial (social shopping and transaction) activities give rise to four distinct states that determine the navigational path of users' online social, and commercial activities. The sequence of actions and transition across the states could show a deterministic pattern in users' navigational behavior. However, given the volume of data and a large number of user actions it is infeasible to model all actions together directly. Thus, we need to summarize the relative importance of actions that users take. In Figure 2, we present the directed graph of two random users' actions across their all sessions. Notice that some of the users never perform some of the actions on this social commerce platform.

---------- *Insert Figure 2 about here* ----------

Methodology

Our modeling approach comprises of three parts. First, we model how to rank the order of the user actions based on their visit. Second, we model the shortest path taken by users to navigate from one state to another. Third, we model the effects of various factors such as ranks, distance, situational variables, and users' characteristics on their visit and time spent in each state.

The specific reasons for these three steps are as follows. First, we note that our dataset is enormous, as more than seventeen thousand sessions are leading to more than ten million actions in merely six months' time by just five hundred users. Therefore, we take advantage of the clickstream data where for each user we have a detailed record of user's actions that lead to their transition from one state to another across various sessions. Such clickstream dataset is suitable for modeling customer using graph-based online attribution modeling [3]. Based on prior studies [1, 38, 59], we summarize the importance of these actions on each state of the social commerce platform using their ranks, which is operationalized based on page rank algorithm [10, 39, 60]. Second, clickstream data capture the specific path taken by the individual user while navigating from one state to another. Previous studies have shown that online behavior is affected by proximity bias [21, 52] or similarity [40, 44]. Therefore, to account for the effect of distance between states, we model the shortest path between states during each session for each user. Third, user engagement in an online environment has been modeled using their incidences and time-spent measures [16, 50]. Therefore, we use an econometric model to capture the effects of various factors on user engagement on such platform using incidence and time-spent.

Ranking of User Actions

Users navigate through different stages of social commerce platform in a typical fashion across their different sessions. However, for a given user a pattern underlines the importance of actions for the individual user. Thus, it is essential to recognize the significance of these actions to optimize users' navigational path and subsequently increase their engagement on the platform [28]. We use the page rank algorithm [10] to find the importance of user actions on such a platform. We note that our calculation of action rank is for each user based on his/her entire sessions.

Let $A = \{a_1, a_2, ..., a_N\}$ be the set of actions available on the platform where N is the total number of actions. Each user h navigates through difference actions, a, for a given session S. We construct a connectivity matrix G_h of size $N \times N$ such that $g_{a_i a_j} = 1$ if the user navigates to action a_i from action a_j, otherwise $g_{a_i a_j} = 0$. We show the construction of matrix G_h in Table 4a-b and Figure 3 using an example of a typical user with six actions (or nodes).

---------- *Insert Table 4a-b and Figure 3 about here* ----------

Let r_{a_i} and c_{a_j} be the row and column sum of matrix G_h such that:

$$r_{a_i} = \sum_j g_{a_i a_j} \tag{1}$$

$$c_{a_j} = \sum_i g_{a_i a_j} \tag{2}$$

Then r_{a_i} and c_{a_j} capture in-degree and out-degree of actions a_i and a_j respectively. Consistent with page rank algorithm, we assume p_h is the probability of user h following a random walk (a typical value of $p = 0.85$ is used by Google in its algorithm and is called damping factor[6]) across these actions. The damping factor p_h is the probability that captures user h's tendency to carry out another action on an online social shopping website. We operationalize p_h as follow:

$$p_h = \frac{n_h}{N} \tag{3}$$

Where n_h is the total number of distinct actions performed by the user h in that session. Let W_h be an $N \times N$ matrix whose elements are as follow:

$$w_{a_i a_j} = \begin{cases} \dfrac{p_h g_{a_i a_j}}{c_{a_j}} + \delta & if \;\; c_{a_j} \neq 0 \\ \dfrac{1}{n} & if \;\; c_{a_j} = 0 \end{cases} \tag{4}$$

Which captures the probability of that action being chosen. W_h is the transition probability matrix of the Markov chain. All the elements of this matrix are positive and less than one, and its column sums are all equal to one. Thus, we can apply a result from matrix theory called Perron-Frobenius theorem that guarantees a nonzero solution to the following equation:

$$x = Wx \tag{5}$$

A solution to this equation exists and is unique to within a scaling factor. If the scaling factor is chosen is such that:

[6] Fu, Lin, and Tsai [29] recommends adoption of damping factor as .85 as a robust and reliable way to determine the importance of web pages.

$$\sum_i x_i = 1 \tag{6}$$

Then, x is the state vector of the Markov chain and is Action Rank. This scaling scheme ensures all elements of x are positive and less than one.

Shortest Distance between Two Actions

User h for given session S navigates through different actions $A = \{a_1, a_2, ..., a_N\}$. We investigate the shortest distance from a source action a_s to a set of all other actions A-a_s. If $e_{uv} = (a_u, a_v)$ is the edge or directed link from action a_u to a_v, then, $w(a_u, a_v)$ or $w(e_{uv})$ is the weight of the links.

The weight of a link represents the cost of visiting from one node to another in that link. We calculate this visit cost based on the total number of clicks required. If a network has N nodes, then the maximum clicks required to visit from one node to another would be N-1 clicks and minimum would be one click. Therefore, we initialize all the weights by maximum clicks required. Next, every time a user navigates from node a_i to a_j the weight on that link $w(a_i, a_j)$ is decreased by one. This decrement of weights can reach all the way to minimum clicks required which is one. Note that one can refine the subsequent decrement of link weight based on other considerations such as a total number of sessions and total distinct nodes (or actions) visited during a session. However, we adopt this algorithm for two reasons, first for its simplicity, and second, it accounts for the frequency of visit between the two nodes. If a user visits a particular node a_j from node a_i more than N-1 times then the weight of that link $w(a_i, a_j)$ becomes one. We outline the algorithm in Table 5. Finally, we use Dijkstra's algorithm [20] to find the shortest path between two nodes. We describe the algorithm in Table 6.

---------- *Insert Table 5 and Table 6 about here* ----------

Econometric Model of Users Engagement

We model user engagement along two dimensions: incidence/visit and time spent in each state.[7] We observe for a given session S whether a user h visits any of the four states C or not through incidence indicator variable $I_{s,c}^{h}$ which takes value 1 if the user visits, otherwise it is 0. The latent utility that the user h gets by this visit is $U_{s,c}^{h}$. Let $T_{s,c}^{h}$ be the time spent by the user in state C during the session S. Note that the variable $T_{s,c}^{h}$ is semi-latent which is partially observed only when the visit takes place otherwise it is 0. We model these two user decisions as follow[8]:

$$U_{s,c}^{h} = \alpha_{c}^{h} + \beta_{c}^{h} AvgRank_{s-1,c^{'}}^{h} + \delta_{c}^{h} SDist_{s-1,c^{'}}^{h} + \chi_{sc}^{h} Situat_{s,c}^{h} + \varepsilon_{s,c}^{h} \tag{7}$$

$$T_{s,c}^{h} = \phi_{c}^{h} + \theta_{c}^{h} AvgRank_{s,c^{'}}^{h} + \lambda_{c}^{h} SDist_{s,c^{'}}^{h} + \gamma_{sc}^{h} Situat_{s,c}^{h} + \xi_{s,c}^{h} \tag{8}$$

Where *AvgRank* is the average rank of all the actions in other states[9], *SDist* is the minimum shortest distance from a particular state C to all other states $c^{'}$ [10]. We note that for incidence model (Equation 7) we use rankings and shortest distances from the previous session $s-1$, whereas for the time-spent model (Equation 8) we use the current session's values. The *Situat* variables capture situations of activities such as time of the day, holidays, and weekend.

The relationship between observed and latent (and semi-latent) variables is as follow:

$$I_{s,c}^{h} = \begin{cases} 1 & if \;\; U_{s,c}^{h} > 0 \\ 0 & otherwise \end{cases} \tag{9}$$

$$t_{s,c}^{h} = \begin{cases} \exp(T_{s,c}^{h}) & if \;\; I_{s,c}^{h} = 1 \\ 0 & otherwise \end{cases} \tag{10}$$

The error terms, $\Sigma = (\varepsilon_{sc}^{h}, \xi_{sc}^{h})^{'}$, are distributed multivariate normal, i.e., $\Sigma \sim MVN(0, \Omega)$. The above formulation gives rise to multivariate type 2 Tobit model.

---

[7] Note that one can also model this at each action level. However, for the sake of parsimony we model at state level.
[8] We note that for the first user session, we operationalize ranking and shortest distance as 0.
[9] Note that using the proposed ranking algorithm the sum of rank across all actions is 1 (Equation 6). Therefore, for model identification we cannot include ranks for all states (or actions). Here, we are interested in how ranks of other nodes (or states) affect visit and time spent on a particular node (or state).
[10] If the visit from a particular state to another is not possible then our algorithm assigns the inter-state distance to *N-1*, where *N* is the total number of possible actions. In such cases, due to multicollinearity we need to drop one variable.

Heterogeneity. We model heterogeneity across users using a hierarchical specification for the users' response parameters. However, for the sake of parsimony, we implement this only for the intercept terms. Thus, let $\mathrm{A}_{hc} = (\alpha_c^h, \phi_c^h)'$, and Z be the matrix of user characteristics, then:

$$\mathrm{A}_{hc} = Z_h \psi_c + \tau_{hc} \tag{11}$$

where $\tau_{hc} \sim MVN(0, \Sigma)$. User characteristics constitute gender and user privacy settings.

Estimation. We use hierarchical Bayesian estimation methods in which multivariate normal distribution is used to model heterogeneity. Markov Chain Monte Carlo (MCMC) techniques are used to simulate the draws for which we use the Gibbs sampler to simulate parameter draws from their posteriors. Furthermore, we use data augmentation techniques for the latent variables to estimate the model parameters. Bayesian methods require priors for which we specify uninformative priors for the model parameters. We use 50,000 iterations for the Markov Chain with the first 40,000 as "burn-in." After ensuring the convergence criterion, we use last 10,000 iterations for the calculation of posterior means and standard deviations of the model parameters for inferences. In Appendix D, we outline the conditionals and data augmentation techniques.

Results

We carry out two analyses at two levels. First, we measure the action ranking and shortest path at the aggregate level. Second, disaggregate level analysis is carried out for the econometric specification of incidence and time spent. We note that for this analysis, actions' rankings and shorted distances between them are calculated for the individual user at the disaggregate level.

Aggregate Level Analysis

The aggregate level analysis reports results from all users' actions across their all sessions. We perform this analysis for ranking of user actions and the shortest distance between the states. This analysis provides general insights into users' path navigation on social commerce platform.

Ranking of User Actions. We present the results of overall ranks of user actions in Table 7. In Figure 4, we show their rankings in a bar graph grouping them together into four distinctly identified states. We find that commercial actions relating to social shopping and transactions are of utmost importance for users on such platforms. Their actions about social activities especially social identification are the lowest rank[11]. Actions pertaining to social interaction matter, but they are ranked lower compared to commercial actions. Thus, based on the social influence theory, we can conclude that social activities supplement users' shopping activities [42, 81].

---------- *Insert Table 7 and Figure 4 about here* ----------

Shortest Path across States. We present the results of the aggregate level shortest path across user states in Table 8. We find users tend to cover longer distances, in terms of a number of clicks required, for social activities than commercial activities. We find that the shortest distance for shopping activities is between 'viewing a product' and 'adding them to cart.' We also observe that after viewing a product either a user adds the product to the cart or makes an entry into his/her favorite list (by adding or removing the item) or sometimes sends the family member an invitation. Thus, users in a shopping state are equally prone to engage either in commercial or social activities. Based on these results, we conclude that social activities lead users to commercial actions due to social engagement [43] and social identity [67].

---------- *Insert Table 8 about here* ----------

Disaggregate Level Analysis

In Table 9 and 10, we present the parameter estimates of the disaggregate level model.

---------- *Insert Table 9 and Table 10 about here* ----------

[11] Note that users accessing the site using tablets or mobile devices need to log in only once.

Social Identification State. Users' visit to social identification state on social commerce platform is positively affected by the previous ranks of other states, with highest impact of transaction state (3.58), followed by shopping state (3.31), and social interaction state (2.48). Thus, users tend to visit and identify in social commerce platform more if their previous activities were related more to transaction and shopping. Consistently, we find that if users' previous distances from social identification state to the transaction (-0.05), and shopping state (-0.79) increase, their probability of visiting the social identification state decreases. However, the distance from social interaction state (1.54) has a positive effect on the visit to social identification state. These results suggest that if users are indulging in social commerce, their cost of traversing the online path on social commerce platform should be reduced, whereas, if they are indulging in entertainment-based activities then having more of these activities increases their probability of a visit to social identification state. Concerning situational factors, we find weekend (0.03) and the gap between sessions (0.01) positively influence visit to social identification whereas evening time (-0.03) has a negative effect.

We find a similar pattern for time spent by users in social identification sate. The current ranks of social interaction state (0.13), shopping state (0.15), and transaction state (0.14) have positive effects on time spent by users on this state. The current distances from social identification state to social interaction state (0.18) and transaction state (0.07) have positive effects on time spent on this state whereas distance to shopping (-0.24) has a negative effect. These results suggest that the users taking shopping-related activities severely minimize their time spent on other activities on the platform. We find that on holidays (0.01) users tend to spend more time, and as the gap between sessions (0.002) increases, the time spent on this state also increases. However, during evening time (-0.01) users tend to spend less time. Concerning user differences, we find male users (-0.11) tend to spend less time whereas users who have given social permission (0.04) and rating permission (0.18) tend to spend more time in this state.

Social Interaction State. Socially entertaining activities such as voting, liking, and commenting are the main activities users can perform in social interaction state. Our empirical results suggest previous ranks of social identification state (1.54) has a positive effect whereas that of shopping (-1.90) and transaction states (-2.00) have negative effects on the visit to this state. Consistently, we find that an increase in previous distances from this state to the social shopping (-0.09) and transaction (-0.49) states have negative effects on a visit to a social interaction state. These results suggest that after commercial activities users tend not to indulge in social or entertainment-based activities. User visit to social interaction states increases during the weekend (0.03) and evening time (0.05) whereas during holidays it goes down (-0.07). User heterogeneity has a significant effect on a visit to this state, with male (-0.19) users visiting less whereas users connected with a Facebook account (1.02), and users who have given permission for social (0.59), rating (1.16) and follower (0.26) tend to visit this state more.

Time spent in this state by the users is only affected by the current ranking of transaction state (-0.21) and has a negative effect. The current distance from this state to a social shopping state (-0.09) has a negative influence on time spent whereas the distances to social identification (0.04) and transaction states (0.07) have positive effects. These results again highlight that the shopping activities before the actual transaction process are a serious consideration for the users that tend to reduce their engagements in other states. Users tend to spend more time on the weekend (0.01) and evening time (0.01) whereas on holidays they spend less time in this state. As the time gap between sessions (-0.001) increases, users spend less time in this state. Concerning user differences, we find that male users (-0.88) spend less time on this state whereas grants with social (0.01), rating (0.99), and follower (0.89) permission have higher tendency to spend time.

Social Shopping State. Users' incidence to social shopping state is negatively affected by the previous ranks of social identification (-2.71) and social interaction (-1.95) states. However, the previous rank of transaction states (1.79) positively affects the current incidence to the social shopping state. Previous distances to social identification (-0.09), social interaction (-0.28) and transaction (-0.08) states have a negative impact on the incidence to this state. These results suggest that users are widely engaged in various social shopping as well as entertainment activities while visiting social shopping states, and they are differentially sensitive to their actions in other states. Also, we find during the weekend (-0.07) and evening time (-0.04) users prefer not to visit this state whereas the gap between sessions (0.01) positively affects their incidence to this state. Grant to rating permission (1.61) has a positive effect on the visit to this state.

We find that there are no effects of current ranks of other states on time spent in this state. However, the current distance from social shopping state to social identification (-0.01) has a negative effect, and to social interaction state (0.004) has a positive effect on the time spent in this state. During weekends (-0.001) users do not spend too much time, however, as the time gap between sessions (0.0002) increases it positively affect their time spent in this state. Male (-0.11) spend less time and users with the Facebook connection (0.06) spend more time in this state.

Transaction State. We find that the incidence to the transaction state is most sensitive to the previous ranks of other states. Specifically, previous ranks of social identification (-0.71) and social interaction (-1.02) have negative effects, whereas, that of social shopping state (0.44) has a positive effect on the visit to this state. For the identification purpose, we have excluded the distance between transaction and identification state as this transition is not observed. We find that the previous distances from transaction to social interaction (-1.82) and social shopping (-0.03) states have negative effects on the visit to this state. Also, we find during the weekend (-0.04) and evening (-0.03) users do not prefer visiting this state, however, as the time gap between sessions (0.01) increases it increases the likelihood of a visit to this state. Concerning user heterogeneity, we find that male users (1.95) tend to visit more than female, and users having a Facebook account (2.08), and who have given a rating (1.66), and follower (1.56) permission tend to visit this state.

Similar to the incidence results we find that time spent by users in this state is negatively affected by current ranks of social identification (-0.27) and social interactions (-0.27) states, whereas that of social shopping state (0.12) has a positive effect. Again, for the identification purpose, we exclude the distance between the transaction and social identification state as this transition is not observed in the entire dataset. The current distances from this state to social interaction (-0.03) and social shopping (-0.03) states have negative effects on the time spent in this state. Users prefer not to spend time during the weekend (-0.02) and evening time (-0.01) in this state. As the time gap between sessions (0.001) increases, the time spent by the users in this state also increases. We find that male users (0.59) spend more time in this state. Furthermore, a user with rating (0.61) and follower (0.44) permission grants tend to spend more time in this state.

General Discussion

We summarize the results of our empirical analysis in Table 11. We gain following insights. First, rank and distance between the states are good predictors of user engagement on the social commerce platform. Second, we find that the effects of rank and distance have a higher impact on users' incidence decision than on their decision to spend time. Third, we observe that there are differential effects across states. In general, ranks of other states have positive effects on visit and time spent in a particular state with 60 percent of the positive parameters. Contrary to this, in general, the distances between states have a negative effect on visit and time spent in a particular state with 60 percent of the negative parameters. Interestingly the effects of inter-state distance on shopping state have all negative effects indicating that shopping is a highly indulgent activity that users perform online, and therefore, should be distraction free. Thus, once users are in shopping state, it is vital that all other states are optimally one click away.

---------- *Insert Table 11 about here* ----------

We find significant effects of many situational variables. On weekends, users prefer to indulge in social activities but not so much in commercial activities. During holidays, both social and commercial activities on such platform are diminished. Finally, as the gap between the sessions increases, user activities in these states tend to increase.

Our results also suggest there is heterogeneity among user activities on social commerce platform. For example, regarding gender differences, we find males have the lesser intensity to indulge in social activities. However, they have a higher propensity to indulge in commercial activities. Furthermore, users connected through social logins (such as Facebook) are socially more active. Moreover, users who have granted social, rating, and follower permission tend to indulge in social activities more actively.

Managerial Implications

The success of social commerce platform depends on user engagement. Our conceptual framework classifies user engagement into social and commercial activities. Furthermore, we empirically capture user indulgence in these activities through visit and time spent on these sites. One of the challenges these firms face is to find the factors that affect user indulgence in these sites. Given the complex network of online social environment and problems associated with capturing user behavior, one can list many explanatory factors that influence user indulgence. This problem necessitates data collection from different sources, which is a daunting task. Our study provides a solution to this problem by utilizing data from a single source, i.e., from the parent social commerce platform. We find that the explanatory factors of our model, rank, distance, situational variables, and user characteristics, have a significant effect on user indulgence captured through visit and time spent. Results from our empirical analysis could provide managers with some guidelines about the design of the social commerce platform. Furthermore, depending on individual user's activities, one can customize the site based on her previous session information to increase the indulgence on the site.

In addition, the objectives of social commerce platforms could be two folds: first, they would like users to visit their platform, and second, they would like users to spend time on their platforms. In our study, we model the former using incidence and latter using time spent. Our results indicate that covariates have differential effects on these two engagement metrics. Therefore, we recommend such platform managers to focus on these engagement metrics individually.

Conclusion

In this study, first, we propose four stages of user engagement on social commerce platforms. Second, using navigational clickstream data from a social commerce platform, we empirically model user engagement along two dimensions: incidence and time spent. We find that the rank and distance between different activities have significant effects on user engagement. Furthermore, situational variables also have significant effects. Users' demographics and online privacy settings have significant explanatory power to capture the heterogeneity across user activities on social commerce platform. We hope the results from this study will provide important insights into the design of social commerce platform and understanding of the online social shopping behavior.

References

1. Abel, M. W., and Shepelyansky, D. L. Google matrix of business process management. *The European Physical Journal B*, 84, 4 (2011), 493-500.
2. Algesheimer, R., Sharad B., Dholakia U.M. and Singh, S.S. The Impact of Customer Community Participation on Customer Behaviors: An Empirical Investigation. *Marketing Science*, 29, 4 (2010), 756-69.
3. Anderl, E., Becker, I., von Wangenheim, F. and Schumann, J. Mapping the customer journey: Lessons learned from graph-based online attribution modeling. International *Journal of Research in Marketing*, 33, 3 (2016), 457-474.
4. Aral, S. and Walker, D. Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Mgt. Science*, 57, 9 (2011), 1623-39.
5. Baghdadi, Y. A framework for social commerce design. *Info. Systems*, 60, (2016), 95-113.
6. Baethge, C., Klier, J. and Klier, M. Social commerce – state-of-the-art and future research directions. *Electronic Markets*, 26, 3 (2016), 269–290.
7. Bai, B., Law, R., and Wen, I. The impact of website quality on customer satisfaction and purchase intentions: Evidence from Chinese online visitors. *International Journal of Hospitality Management*, 27, 3 (2008), 391-402.
8. *BI Intelligence*. The Social Commerce Report: How Facebook, YouTube, Pinterest, and other popular apps are upending the e-commerce space. Jan 25, (2018). (Accessed on Feb 2018 http://www.businessinsider.com/social-commerce-report-2018-1?r=US&IR=T&IR=T)
9. Bilgihan, A., and Bujisic, M. The effect of website features in online relationship marketing: A case of online hotel booking. *Electronic Commerce Research and Applications*, 14, 4 (2015), 222-232.
10. Brin, S. and Page, L. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer networks and ISDN systems*, 31, 1 (1998), 107-17.
11. Bucklin, R.E. and Sismeiro, C. A Model of Web Site Browsing Behavior Estimated on Clickstream Data. *Journal of Marketing Research*, 40, 3 (2003), 249-67.
12. Bucklin, R. E., and Sismeiro, C. Click Here for Internet Insight: Advances in Clickstream Data Analysis in Marketing. *Journal of Interactive Marketing,* 23, 1 (2009), 35-48.
13. *Business News Daily*. Shopping on Social Media: The Future of Social Commerce. October 10, (2017). (Accessed on Feb 2018 https://www.businessnewsdaily.com/6318-future-of-social-commerce.html).
14. Campbell, C., Ferraro, C. and Sands, S. Segmenting consumer reactions to social network marketing. *European Journal of Marketing*, 48, 3-4 (2014), 432-452.
15. Cataldi, M., Di Caro, L. and Schifanella, C. Emerging topic detection on Twitter based on temporal and social terms evaluation. *In Proceedings of the tenth international workshop on multimedia data mining* (p. 4), July (2010). ACM.
16. Chen, CC. Measuring Internet Users' Online Activity: An Application of the Superefficiency Data Envelopment Analysis Model. *Information Society*, 31, 4, (2015), 315-345.
17. Cone (2007). Americans Scrutinize Business Practices in Deciding What to Buy. (Accessed on Feb 2018 http://www.conecomm.com/news-blog/americans-scrutinize-business-practices-in-deciding-what-to-buy).
18. Constantinides, E., Romero, C. L. and Boria, M. A. G. "Social media: a new frontier for retailers?." In European Retail Research pp. 1-28. Wiesbaden *Gabler Verlag*, (2008).
19. Danaher, P.J. Modeling Page Views across Multiple Websites with an Application to Internet

Reach and Frequency Prediction. *Marketing Science*, 26, 3 (2007), 422-437.
20. Dijkstra, E. W. A Note on Two Problems in Connection with Graphs. *Numeriche Mathematik*, 1, (1959), 269-271.
21. Eagle, N., Pentland, A. S. and Lazer, D. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106, 36, (2009), 15274-15278.
22. Edelman, D.C. Branding in the Digital Age. *Harvard Business Review*, 88, 12 (2010), 14-18.
23. *eMarketer*. Social Commerce 2018: Its Influence in the Path to Purchase. December 18, (2017) (Accessed on Feb 2018 https://www.emarketer.com/Report/Social-Commerce-2018-Its-Influence-Path-Purchase/2002175)
24. Fischer, E., and Reuber, A.R. Social Interaction Via New Social Media: (How) Can Interactions on Twitter Affect Effectual Thinking and Behavior?. *Journal of Business Venturing*, 26, 1 (2011), 1-18.
25. Gallino, S. and Moreno, A. Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information," *Management Science*, 60, 6 (2014), 1434-1451.
26. Godes, D. and Mayzlin, D. Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science*, 23, 4 (2004), 545-60.
27. Godes, D., Mayzlin, D., Yubo, C., Sanjiv, D., Dellarocas, C., Pfeiffer, B., Libai, B., Sen, S., Shi, M., and Verlegh, P. The Firm's Management of Social Interactions. *Marketing Letters*, 16, 3-4, (2005), 415-428.
28. Goldfarb, A. State Dependence at Internet Portals. *Journal of Economics & Management Strategy*, 15, 2 (2006), 317-52.
29. Fu, Hwai-Hui, Lin, DKJ, and Tsai, HT. Damping factor in Google page ranking. *Applied Stochastic Models in Business and Industry* 22, 5-6 (2006), 431-444
30. Häubl, G. and Trifts, V. Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids. *Marketing Science*, 19, 1 (2000), 4-21.
31. Hauser, J. R., Urban, G.L. Liberali, G. and Braun, M. Website morphing. *Marketing Science*, 28, 2 (2009), 202-223.
32. Hoffman, D.L., and Fodor, M. Can You Measure the ROI of Your Social Media Marketing. *MIT Sloan Management Review*, 52, 1 (2010), 41-49.
33. Holsing, C. and Schultz, C. D. Empirical insights on the effect of user-generated website features on micro-conversions. *International Journal of E-Business Research*, 9, 4 (2013), 33-46.
34. Huang, Z. and Benyoucef, M. From e-commerce to social commerce: A close look at design features. *Electronic Commerce Research and Applications*, 12, 4 (2013), 246-259.
35. Huang, Z. and Benyoucef, M. User preferences of social features on social commerce websites: An empirical study. *Technological Forecasting and Social Change*, 95, (2015), 57-72.
36. Huang, Z. and Benyoucef, M. The effects of social commerce design on consumer purchase decision-making: An empirical study. *Electronic Commerce Research and Applications*, 25, (2017), 40-58.
37. Johnson, E.J., Moe, W.W., Fader, P.S., Bellman, S. and Lohse, G.L. On the Depth and Dynamics of Online Search Behavior. *Management Science*, 50, 3 (2004), 299-308.
38. Kannan, P.K., and Kopalle, P.K. Dynamic pricing on the Internet: Importance and implications for consumer behavior. *International Journal of Electronic Commerce* 5, 3 (2001): 63-83.
39. Kiss, C. and Bichler, M. Identification of influencers—measuring influence in customer networks. *Decision Support Systems*, 46, 1 (2008), 233-253.

40. Koh, J., Kim, Y. G. and Kim, Y. G. Sense of virtual community: A conceptual framework and empirical validation. *International Journal of Electronic Commerce*, 8, 2, (2003), 75-94.
41. Kozinets, R.V., De Valck, K., Wojnicki, A.C. and Wilner, S. Networked Narratives: Understanding Word-of-Mouth Marketing in Online Communities. *Journal of Marketing,* 74, 2 (2010), 71-89.
42. Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R. and Kannan, P. K. From social to sale: The effects of firm-generated content in social media on customer behavior. *Journal of Marketing*, 80, 1, (2016), 7-25.
43. Lee, D., Hosanagar, K., and Nair, H.S. Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook. *Management Science*, (2018).
44. Lee, M. K., and Turban, E. A trust model for consumer internet shopping. *International Journal of Electronic Commerce*, 6, 1(2001), 75-91.
45. Liang, T.P. and Turban, E. Introduction to the Special Issue Social Commerce: A Research Framework for Social Commerce. *International Journal of Electronic Commerce*, 16, 2 (2011), 5-14.
46. Liang, T. P., Ho, Y. T., Li, Y. W. and Turban, E. What drives social commerce: The role of social support and relationship quality. *International Journal of Electronic Commerce*, 16, 2 (2011), 69-90.
47. Liang, T. P., and Lai H. J. Effect of store design on consumer purchases: an empirical study of on-line bookstores. *Information and Management* 39, 6 (2002), 431-444.
48. Ma, M. and Agarwal, R. Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities. *Information Systems Research*, 18, 1 (2007), 42-67.
49. Mahajan, V. and Venkatesh, R. "Marketing Modeling for E-Business," *International Journal of Research in Marketing*, 17, 2 (2000), 215-25.
50. Mallapragada, G., Chandukala, S. R. and Liu, Q. Exploring the effects of "What"(product) and "Where"(website) characteristics on online shopping behavior. *Journal of Marketing*, 80, 2, (2016), 21-38.
51. McAlexander, J.H., Schouten, J.W. and Koenig, H.F. Building Brand Community. *Journal of Marketing*, 66, 1 (2002), 38-54.
52. Mislove, A., Koppula, H. S., Gummadi, K. P., Druschel, P. and Bhattacharjee, B. Growth of the Flickr social network. *In Proceedings of the first workshop on Online social networks, ACM*, August, (2008), 25-30.
53. Moe, W.W. Buying, Searching, or Browsing: Differentiating between Online Shoppers Using in-Store Navigational Clickstream," *Journal of Consumer Psychology*, 13, 1–2 (2003), 29–39.
54. Montgomery, A.L., Li, S., Srinivasan, K. and Liechty, J.C. Modeling Online Browsing and Path Analysis Using Clickstream Data. *Marketing Science,* 23, 4 (2004), 579-95.
55. Nadeem, W., Andreini, D., Salo, J., and Laukkanen. T. Engaging consumers online through websites and social media: A gender study of Italian Generation Y clothing consumers. *International Journal of Information Management,* 35, 4 (2015), 432-442.
56. Novak, T.P, Hoffman, D.L. and Yung, Y.F. Measuring the Customer Experience in Online Environments: A Structural Modeling Approach. *Marketing Science*, 19, 1 (2000), 22-42.
57. Olbrich, R. and Holsing, C. Modeling consumer purchasing behavior in social shopping communities with clickstream data. *International Journal of Electronic Commerce,* 16, 2 (2011), 15-40.
58. Orlikowski, W. J. Sociomaterial practices: Exploring technology at work. *Organization*

Studies, 28, 9 (2007), 1435-1448.
59. Pagani, M., and Mirabello. A. The influence of personal and social-interactive engagement in social TV web sites. *International Journal of Electronic Commerce*, 16, 2 (2011), 41-68.
60. Page, L., Brin, S., Motwani, R., & Winograd, T. The PageRank citation ranking: Bringing order to the web. Stanford InfoLab. (1999).
61. Park, D. H., Lee, J. and Han, I. The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, 11, 4 (2007), 125-148.
62. Park, Y.H., and Fader, P.S. Modeling Browsing Behavior at Multiple Websites. *Marketing Science*, 23, 4 (2004), 280-303.
63. Pozza, D.I. Multichannel management gets "social." *European Journal of Marketing*, 48, 7-8 (2014), 1274-1295.
64. Reibstein, D.J. What Attracts Customers to Online Stores, and What Keeps Them Coming Back?. *Journal of the Academy of Marketing Science*, 30, 4 (2002), 465-73.
65. Rishika, R., Kumar, A., Janakiraman, R. and Bezawada, R. The Effect of Customers' Social Media Participation on Customer Visit Frequency and Profitability: An Empirical Investigation. *Information Systems Research*, 24, 1 (2013), 108-27.
66. Rutz, O.J. and Bucklin, R.E. Does Banner Advertising Affect Browsing for Brands? Clickstream Choice Model Says Yes, for Some. *Quantitative Marketing and Economics*, 10, 2 (2012), 231-57.
67. Saboo, A. R., Kumar, V., and Ramani, G. Evaluating the impact of social media activities on human brand sales. *International Journal of Research in Marketing*, 33, 3, (2016), 524-541.
68. Schouten, J.W. and McAlexander, J.H. Subcultures of Consumption: An Ethnography of the New Bikers. *Journal of Consumer Research*, 22, 1 (1995), 43-61.
69. Sismeiro, C., and Bucklin, R.E. Modeling Purchase Behavior at an E-Commerce Web Site: A Task-Completion Approach. *Journal of Marketing Research*, 41, 3 (2004), 306-323.
70. Stephen, A. and Toubia, O. Deriving Value from Social Commerce Networks. *Journal of Marketing Research*, 47, 2 (2010), 215-228.
71. Susarla, A, Oh, J.H. and Tan, Y. Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube. *Information Systems Research*, 23, 1 (2012), 23-41.
72. The Guardian (2012), EBay Launches a Mobile Social Shopping Experience as it predicts the busiest online shopping day on Sunday 2 December. Available at: http://internetretailing.net/2012/11/ebay-launches-a-mobile-social-shopping-experience-as-it-predicts-the-busiest-online-shopping-day-on-sunday-2-december/ (Accessed on 28 August 2017).
73. Taylor, S. E., and Thompson, S. C. Stalking the elusive" vividness" effect. *Psychological Review*, 89, 2 (1982), 155.
74. Trusov, M., Bodapati, A. and Bucklin, R.E. Determining Influential Users in Internet Social Networks. *Journal of Marketing Research*, 47, 4 (2010), 643-58.
75. Trusov, M., Bucklin, R.E., and Pauwels, K. Effects of Word-of-Mouth versus Traditional Marketing: Findings from an Internet Social Networking Site. *Journal of Marketing*, 73, 5 (2009), 90-102.
76. Turban, E., Outland, J., King, D., Lee, J. K., Liang, T. P., & Turban, D. C. Electronic Commerce 2018: A Managerial and Social Networks Perspective. (2017). *Springer*.
77. Venkatesh, V. and Agarwal, R. Turning Visitors into Customers: A Usability-Centric Perspective on Purchase Behavior in Electronic Channels. *Management Science*, 52, 3 (2006),

367-82.
78. *Wall Street Journal*. The Shopping Social Network. (2017). (Accessed on Aug 2017 http://www.wsj.com/articles/SB10001424052970204425904578073320375105606).
79. Yadav, M.S., De Valck, K., Hennig-Thurau, T., Hoffman, D.L. and Spann, M. Social commerce: a contingency framework for assessing marketing potential. *Journal of Interactive Marketing*, 27, 4 (2013), 311-323.
80. Yadav, M.S., and Pavlou, P.A. Marketing in computer-mediated environments: Research synthesis and new directions. *Journal of Marketing*, 78, 1 (2014), 20-40.
81. Zhang, X., Li, S., Burke, R. R. and Leykin, A. An examination of social influence on shopper behavior using video tracking data. *Journal of Marketing*, 78, 5, (2014), 24-41.

Table 1: User Activities on Four States of Social Commerce Platform

Action Name	Action Description	Action Code
	Social Activities (Social Identification)	
CreateUser	A new user is created	4
ReturningUser	A registered user has logged in	5
InvCreate	Sent a family member invitation	21
	Social Activities (Social Interaction)	
EntryVote	Voted product	13
RecipeVote	Voted recipe (thumb up, thumb down, foodiemeter)	14
EntryFavourite	Added the favorite product	15
EntryUnfavourite	Removed the favorite product	16
RecipeFavourite	Added favorite recipe	17
RecipeUnfavourite	Removed favorite recipe	18
EntryComment	Commented on product	19
RecipeComment	Commented on recipe	20
	Commercial Activities (Social Shopping)	
EntryView	Viewed a product.	1
EntryToCart	Added product to the shopping cart	2
EntryAdToCart	Added advertised product to the shopping cart	3
RecipeView	Viewed a recipe	6
EntryFromCart	Removed product from the shopping cart	9
EntryAdFromCart	Removed advertised product from shopping cart	10
RecipeToCart	Added recipe to shopping cart	11
RecipeFromCart	Removed recipe from the shopping cart	12
	Commercial Activities (Transaction)	
OrderNew	Started the order process	7
OrderCheckout	Checked out the order, i.e., actually ordered something	8
OrderSelectDeliverySlot	Selected delivery slot for order	22

Table 2: Total Actions Taken by Users on Social Commerce Platform

Action Code	Action Name	Total Count	Mean (action/user)	SD
1	EntryView	9178985	17173.44	12829.58
2	EntryToCart	148946	253.01	239.06
3	EntryAdToCart	12185	38.67	33.04
4	CreateUser	118	1.00	0.00
5	ReturningUser	8381	15.60	30.61
6	RecipeView	987233	1819.64	3995.20
7	OrderNew	10803	21.47	17.75
8	OrderCheckout	5878	11.83	8.40
9	EntryFromCart	73698	122.56	178.75
10	EntryAdFromCart	4802	17.67	19.04
11	RecipeToCart	1492	13.04	18.33
12	RecipeFromCart	832	9.10	14.61
13	EntryVote	960	23.46	116.93
14	RecipeVote	1769	151.51	1260.03
15	EntryFavourite	1280	25.45	32.87
16	EntryUnfavourite	194	6.64	12.01
17	RecipeFavourite	912	9.90	12.34
18	RecipeUnfavourite	135	3.14	4.64
19	EntryComment	8	2.00	2.24
20	RecipeComment	36	2.07	1.83
21	InvCreate	23	2.50	2.14
22	OrderSelectDeliverySlot	5942	11.96	8.50
Total Events		10444612		
Total Sessions		17845		

Table 3: Data Descriptive of Users

Total Users	500	
Total Duration of Activities (in months)	6	
	Mean	**SD**
Sessions per user	31.38	38.45
Actions (per session per user)	9.52	2.03
Time spent per user per session (in minutes)	18.43	10.14
Average activity period per user (in days)	20.68	14.67
	Percentage	
Users connected with the Facebook account	17.40	
Users granting social permission*	76.20	
Users granting rating permission†	95.80	
Users granting follower permission‡	93.40	
Gender (female)	90.00	

*Profile of users with social permission can be viewed publically.
†Users granting rating permission can get their recipes rated by other users.
‡Users granting follower permission can be followed by other users.

Table 4a: An Example of User's Navigation across Actions across Sessions

User ID	Session	Action Sequence
1	1	$1 \to 2 \to 3 \to 4$
1	1	$1 \to 6 \to 1$
1	1	$2 \to 4 \to 1 \to 3$
1	1	$3 \to 6 \to 3$
1	1	$3 \to 4 \to 5 \to 2$

Table 4b: Construction of Matrix G

User		a_i					
	Actions	**1**	**2**	**3**	**4**	**5**	**6**
a_j	**1**	0	0	0	1	0	1
	2	1	0	0	0	1	0
	3	1	1	0	0	0	1
	4	0	1	1	0	0	0
	5	0	0	1	1	0	0
	6	1	0	1	0	0	0

Table 5: Algorithm for Weight of Links

$\forall (a_i, a_j)$	Initialize all link weights to $N-1$
$do\ w(a_i, a_j) = N-1$	
$do\ \ if\ \ a_i \to a_j\ \ is\ \ travered$	If node a_j is traversed from a_i
$then\ \ w(a_i, a_j) = \max(1, w(a_i, a_j) - 1)$	then, decrease the weight all the way to 1 click.

Table 6: Dijkstra's Shortest Path Algorithm

$dist[a_s] = 0$	Distance to source action is set to 0.
$\forall a_i \in (A - a_s)$	All other distances are set to link weight
$do\ dist[a_i] \leftarrow w(a_i, a_j)$	
$S \leftarrow \varnothing$	Set of visited action is set to empty
$Q \leftarrow A$	This queue contains all the actions
$do\ m \leftarrow min\_distance(Q, dist)$	Select action from Q with minimum distance
$S \leftarrow S \bigcup \{u\}$	Add m to the list of visited action
$\forall v \in neighbors[m]$	
$do\ \ if\ \ dist[v] > dist[m] + \mathrm{w(u,m)}$	If new shortest path is found
$then\ \ dist[v] \leftarrow dist[m] + w(u, m)$	Set new value of found shortest path
$return\ \ dist$	Return distance of shortest path

Table 7: Ranks of User Actions on the Social Commerce Platform

Action Description	Action Code	Action Rank
EntryView	1	0.115
EntryToCart	2	0.065
EntryFromCart	9	0.062
RecipeView	6	0.062
OrderNew	7	0.059
OrderCheckout	8	0.052
OrderSelectDeliverySlot	22	0.049
EntryAdToCart	3	0.041
ReturningUser	5	0.040
EntryAdFromCart	10	0.039
RecipeToCart	11	0.038
RecipeVote	14	0.037
RecipeFavourite	17	0.036
RecipeFromCart	12	0.036
EntryVote	13	0.036
EntryFavourite	15	0.035
RecipeUnfavourite	18	0.034
EntryUnfavourite	16	0.034
RecipeComment	20	0.033
EntryComment	19	0.033
InvCreate	21	0.033
CreateUser	4	0.033

Table 8: Shortest Path across States

		TO			
	States	*Social Identification*	*Social Interaction*	*Social Shopping*	*Transaction*
F	*Social Identification*	---	22	1	21
R	*Social Interaction*	21	---	18	21
O	*Shopping*	21	15	---	17
M	*Transaction*	21	20	20	---

Table 9: Parameter Estimates of Visit

	Social Identification		Social Interaction		Social Shopping		Transaction	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Situational Variables								
Weekend	0.0260***	0.0067	0.0281***	0.0101	-0.0698***	0.0093	-0.0563***	0.0067
Holiday	0.0393**	0.0171	-0.0730**	0.0314	-0.0286	0.0236	0.0215	0.0155
Evening Time	-0.0307***	0.0058	0.0525***	0.0095	-0.0401***	0.0091	-0.0266***	0.0052
Days between Sessions	0.0062***	0.0004	-0.0009*	0.0005	0.0077***	0.0011	0.0047***	0.0004
Previous Ranks of States								
Social Identification (SO)			1.5412*	0.9109	-2.7063***	0.9915	-0.7064***	0.0798
Social Interaction (SI)	2.4792***	0.9383			-1.9515*	0.9994	-1.0205***	0.0851
Shopping (SH)	3.3084***	0.7700	-1.9040**	0.7832			0.4430***	0.0632
Transaction (TR)	3.5805***	0.4698	-2.0004***	0.6510	1.7932**	0.7573		
Previous Distance between States								
SO-SI	1.5496***	0.1015						
SO-SH	-0.7897***	0.0442						
SO-TR	-0.0495	0.0628						
SI-SO			0.1056	0.2260				
SI-SH			-0.0903***	0.0195				
SI-TR			-0.4913**	0.2291				
SH-SO					-0.0871**	0.0341		
SH-SI					-0.2785***	0.0375		
SH-TR					-0.0784***	0.0200		
TR-SE								
TR-SI							-1.8176***	0.1222
TR-SH							-0.0263***	0.0028
Intercept	-5.0519***	0.3680	-4.3234***	0.3375	3.9224***	0.3909	4.0011***	0.3297
Hierarchical Parameters								
Gender	-0.3212	0.6599	-0.1898*	0.1002	0.9402	1.2058	1.9468***	0.6712
Facebook Connection	-0.621	0.5138	1.0220***	0.0800	-0.0079	0.9771	2.0836***	0.5483
Social Permission	-0.5644	0.4974	0.5942***	0.0763	0.7635	0.9271	0.2744	0.4962
Rating Permission	0.7731	1.3758	1.1564***	0.2082	1.6138***	0.3305	1.6567*	0.8578
Follower Permission	-0.7927	1.2665	0.2599*	0.1400	1.3928	2.3601	1.5604***	0.4526

*$p \leq 0.10$ (90% CI does not contain 0) **$p \leq 0.05$ (95% CI does not contain 0) ***$p \leq 0.01$ (99% CI does not contain 0) CI-Confidence Interval

Table 10: Parameter Estimates of Time Spent

	Social Identification		Social Interaction		Social Shopping		Transaction	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Situational Variables								
Weekend	0.0019	0.0021	0.0108**	0.0046	-0.0014**	0.0007	-0.0154***	0.0024
Holiday	0.0144***	0.0052	-0.0209	0.0148	-0.0004	0.0026	0.0081	0.0054
Evening Time	-0.0062***	0.0017	0.0130***	0.0044	-0.0026	0.0028	-0.0052***	0.0018
Days between Sessions	0.0015***	0.0002	-0.0008***	0.0003	0.0002**	0.0001	0.0013***	0.0001
Current Ranks of States								
Social Identification (SO)			-0.0351	0.0526	0.0066	0.0115	-0.2695***	0.0546
Social Interaction (SI)	0.1391***	0.0422			0.0058	0.0203	-0.2657***	0.0695
Shopping (SH)	0.1519***	0.0424	0.0674	0.0519			0.1243*	0.0671
Transaction (TR)	0.1393***	0.0412	-0.2070***	0.0527	-0.0029	0.0060		
Current Distance between States								
SO-SI	0.1790***	0.0322						
SO-SH	-0.2442***	0.0351						
SO-TR	0.0680***	0.0258						
SI-SO			0.0391**	0.0198				
SI-SH			-0.0922***	0.0091				
SI-TR			0.0695***	0.0200				
SH-SO					-0.0080***	0.0023		
SH-SI					0.0036*	0.0021		
SH-TR					-0.0009	0.0011		
TR-SO								
TR-SI							-0.0339***	0.0052
TR-SH							-0.0258***	0.0087
Intercept	-0.0351***	0.0039	-0.1651***	0.0178	0.0454***	0.0054	0.7986**	0.3103
Hierarchical Parameters								
Gender	-0.1131***	0.0367	-0.8836***	0.2501	-0.1107*	0.0646	0.5938***	0.0489
Facebook Connection	-0.0248	0.0221	-0.0546	0.1163	0.0553***	0.0199	0.0337	0.0245
Social Permission	0.0362***	0.0087	0.0058***	0.0011	0.0484	0.0570	-0.026	0.0206
Rating Permission	0.1821***	0.0604	0.9932***	0.3668	0.0434	0.2051	0.6083***	0.1158
Follower Permission	0.0650	0.0553	0.8935***	0.3336	0.1694	0.1824	0.4361***	0.0998

*p≤0.10 (90% CI does not contain 0) **p≤0.05 (95% CI does not contain 0) ***p≤0.01 (99% CI does not contain 0) CI-Confidence Interval

Table 11: Summary of Results

			Effects of							
			Social Identification		**Social Interaction**		**Social Shopping**		**Transaction**	
			Rank	*Distance*	*Rank*	*Distance*	*Rank*	*Distance*	*Rank*	*Distance*
Effects on	**Social Identification**	*Visit*			+	+	+	-	+	-
		Time Spent			+	+	+	-	+	+
	Social Interaction	*Visit*	+	+			-	-	-	-
		Time Spent	-	+			+	-	-	+
	Shopping	*Visit*	-	-	-	-			+	-
		Time Spent	NE	-	NE	+			NE	NE
	Transaction	*Visit*	-	NA	-	-	+	-		
		Time Spent	-	NA	-	-	+	-		

NE – Non-significant effect NA– Not Applicable

Figure 1: Conceptual framework of social commerce platform

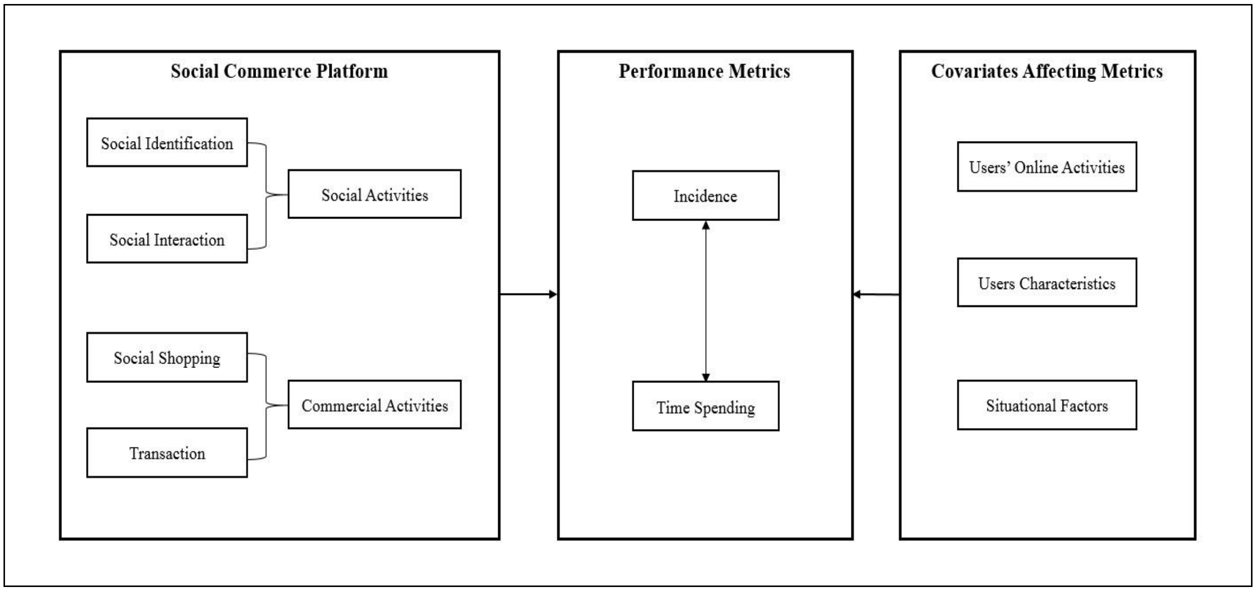
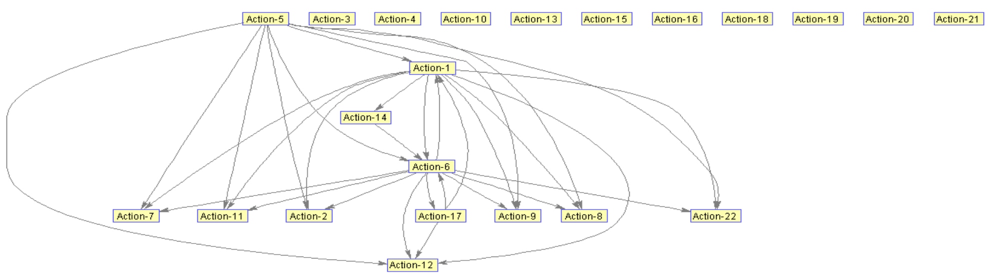


Figure 2: Directed Graphs of Path Navigation of two Users across Sessions

User-1



User-2

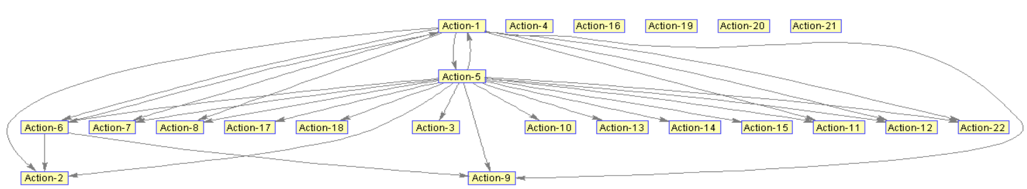


Figure 3: Construction of Graph from the G matrix

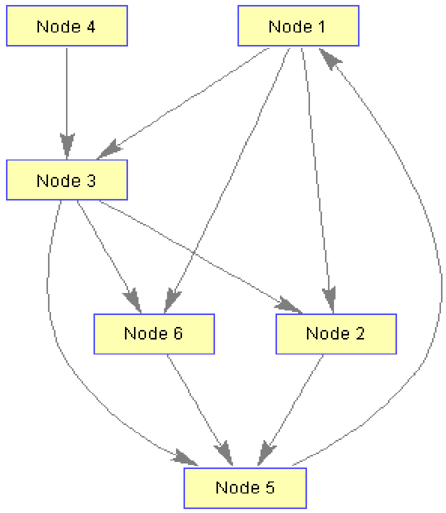


Figure 4: Ranks of User Actions on Social Commerce at Aggregate Level